A Maximum Likelihood Time Delay Estimator in a Multipath Environment Using Importance Sampling

Ahmed Masmoudi, Faouzi Bellili, Sofiène Affes, *Senior Member, IEEE*, and Alex Stéphenne, *Senior Member, IEEE*

***Abstract*—In this paper, we present a new implementation of the maximum likelihood criterion for the estimation of the time delays in a multipath environment and then extend it to the estimation of the time difference of arrival when the transmitted signal is unknown. The new technique implements the concept of importance sampling (IS) to find the global maximum of the compressed likelihood function in a modest computational manner. It thereby avoids traditional complex multidimensional grid search or initialization-dependent iterative methods. Indeed, one of the most interesting features is that it transforms the multi-dimensional search inherent to multipath propagation into a much simpler one-dimensional optimization problem in the delays dimension. Moreover, it guarantees convergence to the global maximum, contrarily to the popular iterative implementation of the maximum likelihood criterion by the well known *expectation maximization* (EM) algorithm. Comparisons with some other methods such as the EM algorithm, MUSIC and accelerated random search (ARS) demonstrates the superiority of the proposed IS-based multipath delay estimator in terms of estimation performance and complexity.**

***Index Terms*—High-resolution methods, iterative methods, importance sampling, maximum likelihood (ML) estimation, Monte-Carlo methods, multipath propagation, timing recovery.**

I. Introduction

TIME delay estimation is a well studied problem with applications in many areas such as radar [2], sonar [3], and wireless communication systems [4]. Typically, to allow estimation of the time delay, an *a priori* known waveform is transmitted through a multipath environment, which consists of several propagation paths, among which the dominant ones, relatively few, are considered. If the transmitted waveform is unknown, only the difference of arrival times can be estimated from the received signals at multiple separated antennas or sensors [5]. In what follows, we will treat the two cases.

These two time delay estimation problems have been extensively studied in previous years [6]–[8]. The maximum likelihood (ML) estimator is well known to be an optimal technique. For the problem at hand, the likelihood function depends on the time delays and on the complex channel coefficients making its solution intractable in a closed form. A direct implementation of this criterion requires a multidimensional grid search, whose complexity increases with the number of unknown delays. Therefore, iterative methods, such as the *expectation maximization* (EM) algorithm, have been developed to achieve the well known Cramér-Rao lower bound (CRLB) at a lower cost. But their performances are closely linked to the initialization values and their convergence may take many complex iterative steps and therefore, a tread-off must be found between complexity and accuracy. Hence, there is yet a need for developing a non-grid-search-based and a non-iterative ML estimator with acceptable complexity. Alternatively, sub-optimal methods based on the eigen-decomposition of the sample covariance matrix, which initially gained much interest in the direction of arrival estimation, were later exploited in the context of time delay estimation [9], [10]. While these suboptimal techniques offer an attractive reduced complexity compared to the grid-search implementation of the ML criterion, they still suffer from heavy computation steps due to the eigenvalue decomposition. Moreover, their performances are relatively poor compared to the ML estimator, especially for closely-spaced delays and/or few numbers of samples.

Motivated by these facts, we derive in this paper a new non-iterative implementation of the ML time delays estimator which avoids the multidimensional grid search by applying:

i) the global maximization theorem of Pincus proposed in [11]
ii) a powerful Monte Carlo technique called importance sampling (IS) offering thereby an efficient tool to find the global maximum of the likelihood function.

Note here that many other traditional Monte Carlo techniques (besides the IS method) can also be successfully applied. However, unlike the IS method, they are usually generated according to a complex probability density function (pdf). Among these methods we cite *pure random search* (PRS) which consists of sampling i.i.d. random variables $\{X_n\}_{n=1}^{R}$ with uniform distribution and then computing $\arg\max(f(X_n), n = 1, \ldots, R)$ where $f(.)$ is the function to maximize. PRS is easy to implement but its convergence is very slow. A modified version of PRS, namely *accelerated random search* (ARS), was proposed in [20] with faster convergence than PRS. Yet the IS technique offers a powerful alternative in which the required realizations are easily generated according to a simpler pdf. Additionally, it offers a more judicious way to process the obtained realizations using a simpler pdf [12]. This method has indeed been applied to the estimation of the direction of arrival (DOA) [14], the joint DOA-Doppler frequency [13], and more recently to the estimation of the time delay in the context of modulated signals and a single propagation path [15].

Manuscript received August 14, 2012; accepted September 18, 2012. Date of publication October 02, 2012; date of current version December 12, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Petr Tichavsky. Work supported by a Canada Research Chair in Wireless Communications and a Discovery Accelerator Supplement from NSERC. This work was published in part at IEEE GLOBECOM, Houston, Texas, December 5–9, 2011 [1].

The authors are with INRS-EMT, Montreal, QC H5A 1K6, Canada (e-mail: masmoudi@emt.inrs.ca; bellili@emt.inrs.ca; affes@emt.inrs.ca; stephenne@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2012.2222402

To compare the current work with that of [15], we can highlight the following differences: The main objective of the work disclosed in [15] was to estimate the delay of a single path from linearly modulated signals in the NDA (non data aided) case. The compressed likelihood function was derived by injecting back the ML estimates of the symbols as function of the unknown single delay. All the derivations were carried out in the time-domain and the problem was inherently one-dimensional. A practical expression of the likelihood function was obtained by using a discrete-time signal model. The problem considered in this paper is to derive maximum likelihood estimators for both the multiple delays and their time differences of arrival in a multipath environment when the transmitted signal is known or unknown, respectively (i.e., active and passive modes, respectively). The optimization problem is therefore multi-dimensional in nature. Hence, we reformulate it in the frequency-domain and the compressed likelihood function is derived by injecting back the ML estimates of the channel multipath coefficients. Our major contribution consists in transforming the intractable multi-dimensional search problem into an equivalent one-dimensional optimization. In fact, we succeed through the manifold-based model to find a *periodogram* that contains all the information about all the unknown delays through its main lobes. Thus, the single- and multipath problems are very different although they both treat time synchronization. We will show that the direct use of the single-path time-domain formulation in [15] will lead to very poor performance compared to our new multipath IS-based estimator, thereby underlining the importance and novelty of the proposed new frequency-domain formulation.

The remainder of this paper is organized as follows. In Section II, we present the system model for the active mode (i.e., known transmitted pulse) and derive the corresponding compressed likelihood function to be maximized. In Section III, we detail the global maximization method applied to our problem. In Section IV, the importance sampling technique is described and then applied to the estimation of the time delays both in active and passive modes. Simulation results are discussed in Section V and, finally, some concluding remarks are drawn out in Section VI.

II. System Model and Compressed Likelihood Function

Consider an *a priori* known signal $x(t)$ transmitted through a multipath environment. The received signal is a superposition of multiple delayed replicas of the known transmitted waveform modelled as follows:

$$y(t) = \sum_{i=1}^{P} \alpha_i x(t-\tau_i) + w(t), \tag{1}$$

where P is the total number of multipath components, $w(t)$ is an additive noise and $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_P]^T$ are the unknown complex path gains resulting from scattering and fading through the propagation medium. In addition, $\{\tau_i\}_{i=1}^{P}$ are the unknown time delays to be estimated and gathered in the vector $\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_P]^T$. If $F_s = \frac{1}{T_s}$ is the sampling frequency, the resulting samples, taken at instances $\{nT_s\}_{n=1}^{N}$ are:

$$y(nT_s) = \sum_{i=1}^{P} \alpha_i x(nT_s-\tau_i) + w(nT_s), \quad n = 0, 1, \ldots, N-1, \tag{2}$$

where N stands for the total number of available samples and the noise $w(nT_s)$ is a complex white Gaussian noise.

In general, the IS principle is suitable for the estimation of non-linear parameters from the general linear models (GLM) described as:

$$\boldsymbol{y} = \boldsymbol{\Phi}(\boldsymbol{\theta})\boldsymbol{s} + \boldsymbol{w}, \tag{3}$$

where $\boldsymbol{y} = [y(0), y(T_s), \ldots, y((N-1)T_s)]^T$ is the received data vector which depends linearly on some nuisance unknown parameters $\boldsymbol{s}$ and non-linearly on the delays $\boldsymbol{\theta}$. However, in contrast to the single-path scenario in [15], the time-domain formulation of the input-output relationship in (2) cannot be used as an input of our estimator without significant loss in performance. Rather, a new frequency-domain formulation will be adopted here. The reasons behind this important claim will be discussed later in Section IV-A once the algorithm will be developed. Here, the received samples are transformed into the frequency-domain where the model could be expressed in a matrix form. In fact, taking the discrete Fourier transform of (2), we obtain:

$$Y(k) = \sum_{i=1}^{P} \alpha_i X(k) e^{-\frac{j2\pi k \tau_i}{N}} + W(k), \quad k = 0, 1, \ldots, N-1, \tag{4}$$

where $\{Y(k)\}_{k=0}^{N-1}$, $\{X(k)\}_{k=0}^{N-1}$ and $\{W(k)\}_{k=0}^{N-1}$ are the discrete Fourier transforms (DFTs) of $y(nT_s), x(nT_s)$ and $w(nT_s)$, respectively. Then, considering this transformation, the channel coefficients vector $\boldsymbol{\alpha}$ and the time delays $\boldsymbol{\tau}$ manifest themselves as the linear and non-linear unknown parameters, respectively. Hence we transform the basic model in (2) into the general form of (3), using a compact representation of (4) as follows:

$$\boldsymbol{Y} = \boldsymbol{\Phi}_a(\boldsymbol{\tau})\boldsymbol{\alpha} + \boldsymbol{W}, \tag{5}$$

in which $\boldsymbol{Y} = [Y(0), Y(1), \ldots, Y(N-1)]^T$ is viewed as the received vector, $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_P]^T$ and $\boldsymbol{W} = [W(0), W(1), \ldots, W(N-1)]^T$ is a complex white Gaussian noise vector containing the DFT coefficients of samples of the additive noise. The matrix[1] $\boldsymbol{\Phi}_a(\boldsymbol{\tau})$ depends only on the unknown delays gathered in the vector $\boldsymbol{\tau}$ and is given by:

$$\boldsymbol{\Phi}_a(\boldsymbol{\tau}) = \left[\boldsymbol{\phi}_a(\tau_1), \boldsymbol{\phi}_a(\tau_2), \ldots, \boldsymbol{\phi}_a(\tau_P)\right], \tag{6}$$

with the columns $\{\boldsymbol{\phi}_a(\tau_i)\}_{i=1}^{P}$ being defined as:

$$\boldsymbol{\phi}_a(\tau_i) = \Big[X(0), X(1)e^{-\frac{j2\pi\tau_i}{N}}, X(2)e^{-\frac{j2\pi 2\tau_i}{N}}, \ldots, X(N-1)e^{-\frac{j2\pi(N-1)\tau_i}{N}}\Big]^T, \tag{7}$$

and $\boldsymbol{X} = [X(0), X(1), \ldots, X(N-1)]^T$ and $\boldsymbol{W} = [W(0), W(1), \ldots, W(N-1)]^T$ are the $(N \times 1)$-dimensional vectors containing the DFT coefficients of samples corresponding to the known transmitted pulse and the additive noise components, respectively.

First, we consider the active model where, in contrast to the passive model treated later in Section IV-B, the transmitted signal $x(t)$ is known to the receiver. Now, following the same

[1]Note that we index $\boldsymbol{\Phi}_a(\boldsymbol{\tau})$ by a to refer to the active mode.

arguments of [16], the likelihood function of the active model (5) is given by:

$$\Lambda(\boldsymbol{\tau},\boldsymbol{\alpha}) \propto p(\boldsymbol{Y};\boldsymbol{\tau},\boldsymbol{\alpha}) = \frac{1}{\pi^N\sigma^{2N}}\exp\Bigg\{-\frac{1}{\sigma^2} \times (\boldsymbol{Y}-\boldsymbol{\Phi}_a(\boldsymbol{\tau})\boldsymbol{\alpha})^H(\boldsymbol{Y}-\boldsymbol{\Phi}_a(\boldsymbol{\tau})\boldsymbol{\alpha})\Bigg\}, \quad (8)$$

where $p(\boldsymbol{Y};\boldsymbol{\tau},\boldsymbol{\alpha})$ is the probability density function (pdf) of $\boldsymbol{Y}$ parameterized by $\boldsymbol{\tau}$ and $\boldsymbol{\alpha}$ and σ^2 is the spectrum power of the noise. Actually, the ML solution $\widehat{\boldsymbol{\tau}}_{ML}$ is defined as the global maximum of the likelihood function in (8) with respect to $\boldsymbol{\tau}$. However, this formulation of the likelihood function imposes a joint estimation of $\boldsymbol{\tau}$ and $\boldsymbol{\alpha}$ which is computationally intensive. Therefore, it is of interest to obtain a likelihood function that depends only on $\boldsymbol{\tau}$ that can be more easily handled. Observing that $\Lambda(\boldsymbol{\tau},\boldsymbol{\alpha})$ is quadratic with respect to $\boldsymbol{\alpha}$, we consider the nuisance parameter, $\boldsymbol{\alpha}$, as deterministic but unknown and substitute, in (8), $\boldsymbol{\alpha}$ by the solution $\widehat{\boldsymbol{\alpha}}(\boldsymbol{\tau})$ which maximizes the log-likelihood function $L(\boldsymbol{\tau},\boldsymbol{\alpha}) = \ln\{\Lambda(\boldsymbol{\tau},\boldsymbol{\alpha})\}$ for a given $\boldsymbol{\tau}$. Indeed, it can be shown that $\widehat{\boldsymbol{\alpha}}(\boldsymbol{\tau})$ is given by:

$$\widehat{\boldsymbol{\alpha}}(\boldsymbol{\tau}) = \left(\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})\right)^{-1}\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{Y}. \quad (9)$$

Replacing $\boldsymbol{\alpha}$ in (8) by $\widehat{\boldsymbol{\alpha}}(\boldsymbol{\tau})$ and omitting the terms that do not interfere in the maximization with respect to $\boldsymbol{\tau}$, we obtain the so-called compressed likelihood function of the system as follows:

$$L_c(\boldsymbol{\tau}) = \frac{1}{\sigma^2}\boldsymbol{Y}^H\boldsymbol{\Phi}_a(\boldsymbol{\tau})\left(\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})\right)^{-1}\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{Y}. \quad (10)$$

III. Global Maximization of the Compressed Likelihood Function

To find the desired ML estimate, we need to maximize the compressed likelihood function in (10) over $\boldsymbol{\tau}$. Yet, $L_c(\boldsymbol{\tau})$ is non-linear with respect to $\boldsymbol{\tau}$; hence, a closed-form solution seems analytically intractable. It is quite common in the current literature to solve this maximization problem in an iterative way, as an alternative for the trivial multidimensional grid search. However, iterative approaches require an initial guess, usually taken from the output of another suboptimal algorithm. The *iterative quadratic* ML (IQML) [17] and the *expectation maximization* (EM) algorithm [6], taken as examples in our simulations, are some of the most famous iterative implementations of the ML estimator. Naturally, the performances of these iterative algorithms depend severely on the available initial guess and may even converge to a local maximum reflecting estimates which are completely different from the real values of the delays (corresponding to the global maximum). Other numerical algorithms have been developed to optimize non-linear functions such as the *simulated annealing* technique [19] and the *accelerated random rearch* (ARS) [20] and could be applied here with various computational costs.

In this context, the global maximization theorem proposed by Pincus [11] offers an alternative to find the global maximum of multidimensional functions, such as the one at hand in (10). Interestingly, it does not require any initialization and guarantees the convergence to the global maximum. The idea is very simple and claims that the solution is given by (11):

$$\widehat{\tau}_i = \lim_{\rho\to\infty}\frac{\int_J\cdots\int_J \tau_i\exp\{\rho L_c(\boldsymbol{\tau})\}\,d\boldsymbol{\tau}}{\int_J\cdots\int_J \exp\{\rho L_c(\overline{\boldsymbol{\tau}})\}\,d\overline{\boldsymbol{\tau}}}, \quad i=1,2,\ldots,P, \quad (11)$$

where J is the interval in which the delays are confined. The only requirements to prove (11) are [11]:

- the function $L_c(.)$ is defined over a compact;
- the function $L_c(.)$ is continuous in this compact;
- the function $L_c(.)$ attains its global maximum at exactly one point of the compact.

Clearly, our objective function (the compressed likelihood function) $\rho L_c(\overline{\boldsymbol{\tau}})$ is continuous and from estimation theory the compressed likelihood function have one global maximum. Therefore convergence is guaranteed. Defining the pseudo-pdf[2] $L'_{c,\rho_0}(\boldsymbol{\tau})$, for some large value of ρ_0, as:

$$L'_{c,\rho_0}(\boldsymbol{\tau}) = \frac{\exp\{\rho_0 L_c(\boldsymbol{\tau})\}}{\int_J\cdots\int_J \exp\{\rho_0 L_c(\overline{\boldsymbol{\tau}})\}\,d\overline{\boldsymbol{\tau}}}, \quad (12)$$

then the optimal value of τ_i is simply given by:

$$\widehat{\tau}_i = \int_J\cdots\int_J \tau_i L'_{c,\rho_0}(\boldsymbol{\tau})d\boldsymbol{\tau}, \quad i=1,2,\ldots,P. \quad (13)$$

Intuitively, we can say that as ρ_0 tends to infinity, the function $L'_{c,\rho_0}(\boldsymbol{\tau})$ becomes a P-dimensional Dirac-delta function centered at the location of the maximum of $L_c(\boldsymbol{\tau})$. Thus, the ML estimate is simply obtained from the evaluation of the P-dimensional integral in (13). Yet, this is a difficult task due to the complexity of the involved integrand function [the pseudo-pdf $L'_{c,\rho_0}(.)$]. One solution is to exploit the fact that $L'_{c,\rho_0}(.)$ is a pseudo-pdf and interpret $\widehat{\tau}_i$ as the expected value of τ_i, the i^{th} element of a vector $\boldsymbol{\tau}$ distributed according to the multidimensional pseudo-pdf $L'_{c,\rho_0}(.)$. Therefore, if one is able to generate R realizations of a random vector, $\{\boldsymbol{\tau}_k\}_{k=1}^R$ according to $L'_{c,\rho_0}(\boldsymbol{\tau})$, it is reasonable to approximate the expected value of $\boldsymbol{\tau}$ using Monte Carlo techniques [12] as follows:

$$\widehat{\boldsymbol{\tau}} = \frac{1}{R}\sum_{k=1}^R \boldsymbol{\tau}_k. \quad (14)$$

Hence, we substitute the complex integration in (13) by a simple samples average. Clearly, as the number of generated values R increases, the variance of the sample mean becomes smaller and $\widehat{\boldsymbol{\tau}}$ gets closer to the global maximum of the compressed likelihood function. Yet a practical issue remains as how to easily generate realizations according to $L'_{c,\rho_0}(\boldsymbol{\tau})$. The proposed pseudo-pdf is a non-linear function of $\boldsymbol{\tau}$ and needs to operate in a multidimensional space, which is not suitable for easy generation of realizations. One solution is to approximate the actual pseudo-pdf by a one-dimensional function and transpose the problem of generating a vector to the generation of P independent variables, then resort to the concept of IS as described in the next section.

[2] $L'_{c,\rho_0}(\boldsymbol{\tau})$ is designated as a pseudo-pdf since it has all the properties of a pdf although $\boldsymbol{\tau}$ is not really a random variable.

IV. The Importance Sampling Based Time Delays Estimation

A. IS Concept

Importance sampling [18] is a Monte Carlo technique which makes use of an alternative distribution (carefully designed) to generate realizations. It is usually applied when the original distribution does not have a practical form, like $L'_{c,\rho_0}(.)$ given in our problem.

The approach is based on the following simple observation on the integral involved in (13):

$$\int_J \cdots \int_J \boldsymbol{\tau} L'_{c,\rho_0}(\boldsymbol{\tau}) d\boldsymbol{\tau} = \int_J \cdots \int_J \boldsymbol{\tau} \frac{L'_{c,\rho_0}(\boldsymbol{\tau})}{g'(\boldsymbol{\tau})} g'(\boldsymbol{\tau}) d\boldsymbol{\tau}, \quad (15)$$

where $g(\boldsymbol{\tau})'$ is also assumed to have all the properties of a pdf, called normalized importance function (IF). Then, the right-hand side of (15) is interpreted as the mean of $\tau_i \frac{L'_{c,\rho_0}(\boldsymbol{\tau})}{g'(\boldsymbol{\tau})}$ when $\boldsymbol{\tau}$ is generated according to $g'(\boldsymbol{\tau})$. Unlike $L'_{c,\rho_0}(.)$, it is of interest to choose $g(\boldsymbol{\tau})'$ to be a simple function of $\boldsymbol{\tau}$. Then, we use Monte-Carlo approximation to numerically compute the expectation as done in (14):

$$\int_J \cdots \int_J \boldsymbol{\tau} \frac{L'_{c,\rho_0}(\boldsymbol{\tau})}{g'(\boldsymbol{\tau})} g'(\boldsymbol{\tau}) d\boldsymbol{\tau} \approx \frac{1}{R} \sum_{k=1}^{R} \boldsymbol{\tau}_k \frac{L'_{c,\rho_0}(\boldsymbol{\tau}_k)}{g'(\boldsymbol{\tau}_k)}, \quad (16)$$

where $\boldsymbol{\tau}_k$ is now the k^{th} realization of $\boldsymbol{\tau}$ according to $g'(.)$.

Clearly, the choice of $g'(.)$ affects the estimation performance. An inappropriate choice of $g'(.)$ may need a large number of realizations R to reduce the estimation variance and result in a higher computational complexity. Therefore, the value of R depends on how much $g'(.)$ resembles $L'_{c,\rho_0}(.)$. In the ideal case, generations according to $g(.)'$ are the same as if they were generated according to $L'_{c,\rho_0}(.)$. Therefore, ideally, the shapes of the two functions $g'(.)$ and $L'_{c,\rho_0}(.)$ should be similar to reduce the variance of the estimator given by (16) [14]. On the other hand, we should keep in mind that $g'(.)$ has to be simple enough so that realizations can be easily generated. Thus some tradeoffs are required to choose a function as simple as possible yet similar to $L'_{c,\rho_0}(.)$. In what follows, we will show that owing some simplifications of $L'_{c,\rho_0}(.)$, we can build an appropriate function $g'(.)$ to properly generate variables.

Now, coming back to the expression of the actual compressed likelihood function in (10), the inverse matrix $(\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau}))^{-1}$ makes the compressed likelihood function, and consequently the pseudo-pdf $L'_{c,\rho_0}(.)$, very non-linear with respect to $\boldsymbol{\tau}$. One can approximate $\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})$ by a diagonal matrix to avoid a heavy computation of the inverse. In fact, the diagonal elements of $\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})$ are given by:

$$\left[\left(\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})\right)\right]_{l;l} = \sum_{k=0}^{N-1} |X(k)|^2, \quad l = 1, 2, \ldots, P, \quad (17)$$

and its off-diagonal elements are:

$$\left[\left(\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})\right)\right]_{m;n} = \left|\sum_{k=0}^{N-1} |X(k)|^2 \exp\left\{\frac{j2\pi k(\tau_m - \tau_n)}{N}\right\}\right|, \quad m, n = 1, 2, \ldots, P, m \neq n. \quad (18)$$

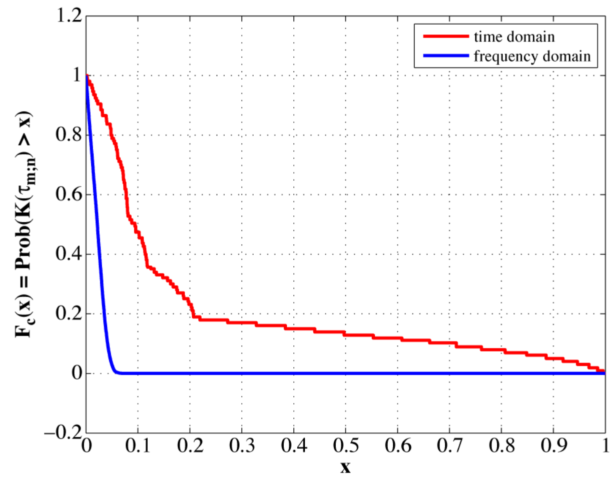


Fig. 1. Complementary cumulative distribution function of the ratio $K(\tau_{m;n})$ with a chirp transmitted signal.

It is easy to verify, for $\tau_m \neq \tau_n$, that:

$$\left[\left(\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})\right)\right]_{m;n} < \left[\left(\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})\right)\right]_{l;l}. \quad (19)$$

This inequality gives a week proof to approximate $\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})$ by a diagonal matrix. We verify statistically that $[(\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau}))]_{m;n}$ is much smaller than $[(\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau}))]_{l;l}$ with very high probability for almost all possible values of the delay difference $\tau_m - \tau_n$. To that end, we consider $\tau_{m;n} = \tau_m - \tau_n$ as a random variable uniformly distributed in[3] $[-T, T]$ and we define $K(\tau_{m;n})$, the ratio (18)/(17), as follows:

$$K(\tau_{m;n}) = \frac{\left|\sum_{k=0}^{N-1} |X(k)|^2 \exp\left\{\frac{j2\pi k(\tau_m - \tau_n)}{N}\right\}\right|}{\sum_{k=0}^{N-1} |X(k)|^2}. \quad (20)$$

Then, we plot in Fig. 1 (T is equal to $\frac{NT_s}{2}$ there) the complementary cumulative distribution function of $K(\tau_{m;n})$ (randomized according to $\tau_{m;n}$), to verify that the diagonal elements of $\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})$ are indeed dominant, with very high probability, compared to its off-diagonal elements. Therefore, we adopt the following well-justified approximation:

$$\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau}) \approx \left(\sum_{k=0}^{N-1} |X(k)|^2\right) \boldsymbol{I}_p, \quad (21)$$

where $\boldsymbol{I}_p$ is the $p \times p$ identity matrix. Then, we define the importance function, $g_{\rho_1}(.)$, in the active case as:

$$g_{\rho_1}(\boldsymbol{\tau}) = \exp\left\{\frac{\rho_1}{\sigma^2 \sum_{k=0}^{N-1} |X(k)|^2} \boldsymbol{Y}^H \boldsymbol{\Phi}_a(\boldsymbol{\tau}) \boldsymbol{\Phi}_a^H(\boldsymbol{\tau}) \boldsymbol{Y}\right\}, \quad (22)$$

where ρ_1 is another constant different from ρ_0 for some practical reasons. A further discussion on the appropriate choice of ρ_0 and ρ_1 is left to the end of this section.

After some easy algebraic manipulations, we express (22) as:

$$g_{\rho_1}(\boldsymbol{\tau}) = \prod_{i=1}^{P} \exp\left\{\frac{\rho_1}{\sigma^2 \sum_{k=0}^{N-1} |X(k)|^2} I(\tau_i)\right\}, \quad (23)$$

[3] We consider here that the delays do not exceed T.

where $I(\tau_i)$ is the periodogram of the data in the frequency-domain evaluated at each delay τ_i as follows:

$$I(\tau_i) = \left| \sum_{k=0}^{N-1} X(k)^* Y(k) \exp\left\{ \frac{j2\pi(k-1)\tau_i}{N} \right\} \right|^2. \quad (24)$$

Note here that the multiplicative terms $X(k), k = 0, \ldots, N-1$ act as weighting factors. In fact, $Y(k)$ is the DFT coefficient of the received signal. Outside the transmitted signal frequency band, $Y(k)$ is composed only of noise. In this band, $X(k)$ is almost equal to zero. Therefore the product $X(k)Y(k)$ is close to zero and this reduces the effect of noise in $I(.)$. This property improves considerably the performance of the estimator compared to some other approaches where the received signal is divided, in the frequency-domain, by the DFT of the known transmitted waveform [9]. Actually, this operation is not suitable for narrowband signals since it results in some harmful effects by amplifying the additive noise in the low-energy frequencies. It is suitable only for wideband signals, in contrast to our algorithm which is also well adapted to narrowband signals.

Now, we comment on the advantage of this choice in (23) for the importance function (IF). First, we notice that the joint contribution of the different delays in $g_{\rho_1}(.)$ is separable into the product of their individual contributions as seen from (23). Therefore, we successfully transform the multi-dimensional search problem into a one-dimensional optimization. Hence, we substitute the brute generation of realizations of the vector $\boldsymbol{\tau}$ according to a multi-dimensional pdf to the generation of P independent scalar realizations (i.e., one realization for each entry of $\boldsymbol{\tau}$) using the elementary IF, $\overline{g}_{\rho_1}(\tau_i)$, defined as:

$$\overline{g}_{\rho_1}(\tau_i) = \frac{\exp\left\{ \frac{\rho_1}{\sigma^2 \sum_{k=0}^{N-1} |X(k)|^2} I(\tau_i) \right\}}{\int_J \exp\left\{ \frac{\rho_1}{\sigma^2 \sum_{k=0}^{N-1} |X(k)|^2} I(\overline{\tau}) \right\} d\overline{\tau}}. \quad (25)$$

Finally, the normalized IF is given by:

$$g'_{\rho'_1}(\boldsymbol{\tau}) = \frac{\prod_{i=1}^{P} \exp\left\{ \rho'_1 I(\tau_i) \right\}}{\int_J \cdots \int_J \prod_{i=1}^{P} \exp\left\{ \rho'_1 I(u_i) \right\} du_i}, \quad (26)$$

with

$$\rho'_1 = \frac{\rho_1}{\sigma^2 \sum_{k=0}^{N-1} |X(k)|^2}. \quad (27)$$

We mention that the choices of the parameters ρ'_1 and ρ_0 are of great importance since they affect the performance of the new estimator. In fact, as already mentioned, $g'_{\rho'_1}(\boldsymbol{\tau})$ is separable as the product of P elementary IFs, $\overline{g}_{\rho'_1}(.)$, corresponding to each delay τ_i (i.e., $g'_{\rho'_1}(\boldsymbol{\tau}) = \prod_{i=1}^{P} \overline{g}_{\rho'_1}(\tau_i)$). Hence, in practice, we use the same $\overline{g}_{\rho'_1}(\tau)$ P times to generate the P elements of the vector $\boldsymbol{\tau}$. Actually, for a noise-free observation, the function $\overline{g}_{\rho'_1}(.)$ exhibits exactly P lobes centered at the locations of the true delays and at each run, a realization is generated from the vicinity of one of the P lobes. However, in the presence of additive noise, other secondary lobes appear and ultimately affect the generated values. For this reason, the parameter ρ'_1 should be increased to render the objective function $\overline{g}_{\rho'_1}(.)$ more peaked around the actual delays $\{\tau_i\}_{i=1}^{P}$. This behavior is illustrated in Fig. 2 where we plot the function $\overline{g}_{\rho'_1}(.)$ for two values of ρ'_1.

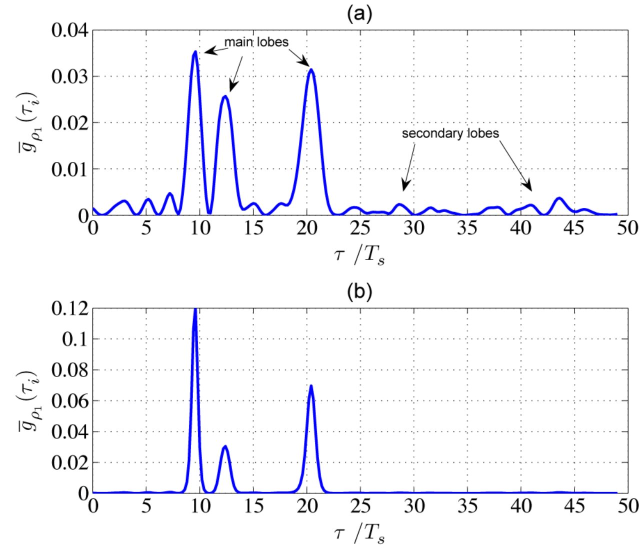


Fig. 2. Plot of $\overline{g}_{\rho'_1}(.)$ at $\mathrm{SNR} = 10$ for (a) $\rho'_1 = 1$ and (b) $\rho'_1 = 6$.

Yet, we observe that ρ'_1 cannot be increased indefinitely. In fact, very large values of ρ'_1 will ultimately destroy some useful lobes and hence useful realizations may not be generated. Obviously, proper choice of ρ'_1 is of great importance. Its optimal value should be the highest one that makes at least[4] P main lobes[5] appear in $\overline{g}_{\rho'_1}(.)$. Moreover, by attenuating the secondary lobes, we reduce the probability of generating undesired realizations. Consequently a good choice of ρ'_1 reduces the number of necessary realizations R and hence the complexity of the estimator.

Recall that the normalized IF $g'_{\rho'_1}(\boldsymbol{\tau})$ is built upon an approximation of the actual compressed likelihood function which results in biased estimates of the delays, especially at low SNR values. However, we emphasize here the fact that this bias can be reduced by the presence of the actual compressed likelihood function in the weighting factor $\frac{L'_{c,\rho_0}(\boldsymbol{\tau})}{g'_{\rho'_1}(\boldsymbol{\tau})}$ in (16). Thus, we can maximize the contribution of $L'_{c,\rho_0}(.)$ in the weighting factor by choosing ρ'_1 smaller than ρ_0. The IS-based estimator requires the generation of realizations according to $\overline{g}_{\rho'_1}(.)$ then evaluating the following mean values:

$$\hat{\tau}_i = \frac{1}{R} \sum_{k=1}^{R} \boldsymbol{\tau}_k(i) \frac{L'_{c,\rho_0}(\boldsymbol{\tau}_k)}{g'_{\rho'_1}(\boldsymbol{\tau}_k)}, \quad (28)$$

where $\boldsymbol{\tau}_k$ is the k^{th} generated vector and $\boldsymbol{\tau}_k(i)$ refers to its i^{th} element. The delays can actually take any positive value, but in practice, they are confined in the interval $[0, T]$ where T is any positive real value that can be chosen high enough[6] so that $\tau_i \in [0, T]$ for $i = 1, 2, \ldots, P$. Therefore, since the parameters are bounded from below and above, it is more convenient to use the circular mean [21] instead of the linear mean in (28). In this

[4]$\overline{g}_{\rho'_1}(.)$ should have exactly P lobes, but the additive noise makes other relatively small secondary lobes appear.

[5]Main lobes are the ones which are centered around the actual unknown delays while secondary lobes are those which are centered around value that do not correspond to any one of the true delays.

[6]In network communications, the delays are usually confined in the symbol duration, whereas for radar and sonar systems, the symbol duration does not really exist and the observation window must be longer than the largest delay.

context, the alternative formulation of the IS-based estimator is given by:

$$\widehat{\tau}_i = \frac{1}{2\pi T}\angle \frac{1}{R}\sum_{k=1}^{R} F(\boldsymbol{\tau}_k)\exp\left\{j2\pi\frac{\boldsymbol{\tau}_k(i)}{T}\right\}, \tag{29}$$

where $F(\boldsymbol{\tau}_k)$ is the weighting factor defined by:

$$F(\boldsymbol{\tau}_k) = \frac{L'_{c,\rho_0}(\boldsymbol{\tau}_k)}{g'_{\rho'_1}(\boldsymbol{\tau}_k)}. \tag{30}$$

From the formulation in (29), we only need to find the angle of a complex number. Therefore, any positive multiplicative term will not affect the final result. Thus, the two strictly positive constants $\int_J \cdots \int_J \exp\{\rho_0 L_c(\boldsymbol{x})\}d\boldsymbol{x}$ and $\int_J \cdots \int_J \prod_{i=1}^{P}\exp\{\rho'_1 I(u_i)\}du_i$, used in the normalization of $L'_{c,\rho_0}(.)$ and $g'_{\rho'_1}(\boldsymbol{x})$, respectively, can be dropped. However, the exponential terms in both the numerator and the denominator of the weighting factor $F(.)$ may result in an overflow in the computation. To circumvent this problem, $F(.)$ is substituted by $F'(.)$:

$$F'(\boldsymbol{\tau}_k) = \exp\left\{\rho_0 L_c(\boldsymbol{\tau}_k) - \rho'_1\sum_{i=1}^{P} I\left(\tau_k(i)\right) - \max_{1\leq l\leq R}\left(\rho_0 L_c(\boldsymbol{\tau}_l) - \rho'_1\sum_{i=1}^{P} I\left(\boldsymbol{\tau}_l(i)\right)\right)\right\}. \tag{31}$$

Note from (31) that the arguments of the exponential terms are either negative or zero and that the values of the exponential cannot exceed one.

Summary of Steps: In the following, we recapitulate the different steps for the direct implementation of the new algorithm:

1) Compute the DFT $[Y(0), Y(1), \ldots, Y(N-1)]$ of the received signal samples.
2) Use the Fourier transform coefficients to evaluate the periodogram according to (24).
3) Compute the samples of the one-dimensional pdf $\overline{g}_{\rho'_1}(.)$, used for the generation of the required realizations, at K points as:

$$\overline{g}_{\rho'_1}(\tau_l) = \frac{\exp\{\rho'_1 I(\tau_l)\}}{\sum_{k=1}^{K}\exp\{\rho'_1 I(\tau_k)\}}, \quad l = 1, 2, \ldots, K, \tag{32}$$

where K is the total number of points[7] in the interval J, i.e., $J = \left\{0, \frac{1}{K}, \ldots, \tau_{max} - \frac{1}{K}, \tau_{max}\right\}$ which determine the precision of tuning. Note that we substitute the integration in the denominator of $\overline{g}_{\rho'_1}(.)$ by a summation over the discrete points of the interval J.

4) Generate one realization of $\boldsymbol{\tau}$ using $g'_{\rho'_1}(.)$. To do so, we generate realizations according to $\overline{g}_{\rho'_1}(.)$ P times to retrieve one realization of the P-dimensional vector $\boldsymbol{\tau}$. More details on this point are left to the Appendix. Repeat this step $R-1$ times.
5) Evaluate the weighting factor $F'(\boldsymbol{\tau}_i)$ for $i = 1, 2, \ldots, R$ and compute the circular mean of the generated values balanced by the weighting factors to find the ML estimate of the multiple unknown delays. Note that we must evaluate the term $\rho_0 L_c(\boldsymbol{\tau}_l) - \rho'_1\sum_{m=1}^{P} I(\boldsymbol{\tau}_l(m))$ for all generated vectors $\{\boldsymbol{\tau}_l\}_{l=1}^{R}$ before computing $F'(\boldsymbol{\tau}_i)$.

Frequency—vs. Time-Domain Formulation: To better justify our frequency-domain formulation, considering the time-domain formulation in (2), a compact representation of the resulting samples is given by:

$$\boldsymbol{y} = \boldsymbol{\Phi}_t(\boldsymbol{\tau})\boldsymbol{\alpha} + \boldsymbol{w}, \tag{33}$$

in which $\boldsymbol{y} = [y(0), y(T_s), \ldots, y((N-1)T_s)]$, $\boldsymbol{w} = [w(0), w(T_s), \ldots, w((N-1)T_s)]$ and $\boldsymbol{\Phi}_t(\boldsymbol{\tau})$ depends on the time delays vector and is given by:

$$\boldsymbol{\Phi}_t(\boldsymbol{\tau}) = [\boldsymbol{\phi}_t(\tau_1), \boldsymbol{\phi}_t(\tau_2), \ldots, \boldsymbol{\phi}_t(\tau_P)], \tag{34}$$

with the columns $\{\boldsymbol{\phi}_t(\tau_i)\}_{i=1}^{P}$ being defined as:

$$\boldsymbol{\phi}_t(\tau_i) = [x(-\tau_i), x(T_s - \tau_i), \ldots, x((N-1)T_s - \tau_i)]^T. \tag{35}$$

However, unlike the proposed formulation in the frequency-domain, the problem with this time-domain representation is that the approximation

$$\boldsymbol{\Phi}_t^H(\boldsymbol{\tau})\boldsymbol{\Phi}_t(\boldsymbol{\tau}) \approx \boldsymbol{I}_P, \tag{36}$$

would no longer be accurate. In fact, following the same steps from (17)to (20), the diagonal elements of $\boldsymbol{\Phi}_t^H(\boldsymbol{\tau})\boldsymbol{\Phi}_t(\boldsymbol{\tau})$ will be given by:

$$\left[\left(\boldsymbol{\Phi}_t^H(\boldsymbol{\tau})\boldsymbol{\Phi}_t(\boldsymbol{\tau})\right)\right]_{l;l} = \sum_{k=0}^{N-1}\left|x(kT_s - \tau_l)\right|^2, \quad l = 1, 2, \ldots, P, \tag{37}$$

and its off-diagonal elements are:

$$\left[\left(\boldsymbol{\Phi}_t^H(\boldsymbol{\tau})\boldsymbol{\Phi}_t(\boldsymbol{\tau})\right)\right]_{m;n} = \left|\sum_{k=0}^{N-1} x^*(kT_s - \tau_m)x(kT_s - \tau_n)\right|, \quad m, n = 1, \ldots, P, m \neq n. \tag{38}$$

Similarly, we can consider here the delay difference $\tau_{m,n} = \tau_m - \tau_n$ as a random variable and define, as done in the frequency-domain, $K(\tau_{m;n})$, the ratio (38)/(37), as follows:

$$K(\tau_{m;n}) = \frac{\left|\sum_{k=0}^{N-1} x^*(kT_s - \tau_m)x(kT_s - \tau_n)\right|}{\sum_{k=0}^{N-1}\left|x(kT_s - \tau_m)\right|^2}. \tag{39}$$

In Fig. 1, we also plot the complementary cumulative distribution function of $K(\tau_{m;n})$ (randomized according to $\tau_{m;n}$) to compare this time-domain ratio with the one earlier obtained in the frequency-domain. Clearly, the probability that the diagonal elements of $\boldsymbol{\Phi}_t^H(\boldsymbol{\tau})\boldsymbol{\Phi}_t(\boldsymbol{\tau})$ are dominant in the time-domain is significantly lower than the probability that the diagonal elements of $\boldsymbol{\Phi}_a^H(\boldsymbol{\tau})\boldsymbol{\Phi}_a(\boldsymbol{\tau})$ are dominant in the frequency-domain.

[7] K stands for the number of points $\{\tau_l\}_{l=0}^{K-1}$ for which we evaluate the one-dimensional pseudo-pdf $g'_{\rho'_1}(.)$ and whose obtained values are used during the generation of the corresponding random realizations. In fact, the interval $J = [0, \tau_{max}]$, where τ_{max} is the maximum delay value, is discretized into small equal-size bins of width $\frac{\tau_{max}}{K}$ over which the points $\tau_l = \frac{l}{K}\tau_{max}$ for $l = 0, 1, \ldots, K-1$ are considered. It is clear, then, that there is degree of freedom regarding the choice of K. For instance, K can be freely selected to meet any intended resolution and/or accuracy. Actually, if we aim at estimating the unknown delays with a resolution of $\Delta\tau^{(res)}$, then we should simply select K to satisfy $\frac{\tau_{max}}{K} \leq \Delta\tau^{(res)}$ or $K \geq \frac{\tau_{max}}{\Delta\tau^{(res)}}$. Of course, as K increases, the computational cost becomes higher.

Therefore, we would not be able to neglect the off-diagonal elements of $\boldsymbol{\Phi}_t^H(\boldsymbol{\tau})\boldsymbol{\Phi}_t(\boldsymbol{\tau})$ compared to its diagonal elements and thus a diagonal approximation of $\boldsymbol{\Phi}_t^H(\boldsymbol{\tau})\boldsymbol{\Phi}_t(\boldsymbol{\tau})$ would no longer be accurate and hence possible in the time-domain. We recall here that this approximation is extremely important in order to obtain a simple importance function as the one obtained in the frequency-domain.

To conclude, unlike the time-domain formulation in [15], proposed developments in the frequency-domain allow us to obtain a very simple one-dimensional importance function, one of the key contributions of this work. The net advantages of the new frequency-domain formulation against the time-domain formulation in [15] will be clearly illustrated by simulations later in Section V.

B. Time Delays Estimation in Passive Systems

In a passive system, the transmitted signal is considered to be unknown. In this case, only the time difference of arrival (TDOA) can be estimated from multiple received signals at spatially separated destinations [5]. In this section, we assume, without loss of generality, the presence of two separated sensors. The received signals at these two sensors are modelled as:

$$y_1(t) = \sum_{i=1}^{P_1} \alpha_{1;i} x(t-\tau_{1;i}) + w_1(t), \tag{40}$$

$$y_2(t) = \sum_{i=1}^{P_2} \alpha_{2;i} x(t-\tau_{2;i}) + w_2(t), \tag{41}$$

where $\{\tau_{n;i}\}_{i=1}^{P_n}$ and $\{\alpha_{n;i}\}_{i=1}^{P_n}$, for $n=1,2$, are the delays and the complex gains of the received signal at the n^{th} sensor and $\{P_n\}_{n=1}^{2}$ are the known numbers of multipath components. For the sake of simplicity, suppose that $y_1(t)$ has only one signal component ($P_1 = 1$). The received signal at this sensor is considered as a reference and hence it is assimilated to a noisy known signal. Then, similarly to (4), we express the sampled signals (40) and (41) in the frequency-domain as:

$$Y_1(k) = \alpha_{1;1} X(k) \exp\left\{-\frac{j2\pi k \tau_{1;1}}{N}\right\} + W_1(k), \tag{42}$$

$$\begin{aligned} Y_2(k) &= \sum_{i=1}^{P_2} \alpha_{2;i} X(k) \exp\left\{-\frac{j2\pi k \tau_{2;i}}{N}\right\} + W_2(k), \\ & k = 0,1,\ldots,N-1, \end{aligned} \tag{43}$$

where $\{Y_1(k)\}_{k=0}^{N}$, $\{Y_2(k)\}_{k=0}^{N}$, $\{W_1(k)\}_{k=0}^{N}$ and $\{W_2(k)\}_{k=0}^{N}$ are N samples of the Fourier transform of samples of $y_1(t)$, $y_2(t)$, $w_1(t)$ and $w_2(t)$, respectively. As mentioned above, the TDOAs will be estimated by considering the received signal in the first sensor as a reference. This simplifies to the estimation of the P_2 delay differences $\Delta_\tau^{(i)} = \tau_{2;i} - \tau_{1;1}$ for $i = 1,3,\ldots,P_2$. Therefore, we rewrite (43) as follows:

$$Y_2(k) = \sum_{i=1}^{P_2} \beta_i Y_1(k) \exp\left\{-\frac{j2\pi k \Delta_\tau^{(i)}}{N}\right\} + W_p(k), \tag{44}$$

in which

$$\beta_i = \frac{\alpha_{2;i}}{\alpha_{1;1}}, \quad i = 1,2,\ldots,P_2, \tag{45}$$

$$W_p(k) = W_2(k) - \sum_{i=1}^{P_2} \beta_i W_1(k) \exp\left\{-\frac{j2\pi k \Delta_\tau^{(i)}}{N}\right\}. \tag{46}$$

Doing so, we highlight the parameters of interest in the expression of $Y_2(k)$. Moreover, there is an analogy between the formulation of the active case in (4) and the passive one in (44). More precisely, the major difference is in the reference signal ($\boldsymbol{X}$ for the active case and $\boldsymbol{Y}_1$ for the passive case).

Then, gathering all the frequency samples, we obtain the following matrix representation:

$$\begin{aligned} \boldsymbol{Y}_2 &= [Y_2(0), Y_2(1), \ldots, Y_2(N-1)]^T \\ &= \boldsymbol{\Phi}_p(\boldsymbol{\Delta}_\tau)\boldsymbol{\beta} + \boldsymbol{W}_p, \end{aligned} \tag{47}$$

where the matrix $\boldsymbol{\Phi}_p(\boldsymbol{\Delta}_\tau)$ is function of the TDOAs defined as:

$$\boldsymbol{\Phi}_p(\boldsymbol{\Delta}_\tau) = \left[\boldsymbol{\phi}_p\left(\Delta_\tau^{(1)}\right), \boldsymbol{\phi}_p\left(\Delta_\tau^{(2)}\right), \ldots, \boldsymbol{\phi}_p\left(\Delta_\tau^{(P_2)}\right)\right], \tag{48}$$

$$\begin{aligned} \boldsymbol{\phi}_p\left(\Delta_\tau^{(i)}\right) = \Bigg[& Y_1(0), Y_1(1)\exp\left\{-\frac{j2\pi\Delta_\tau^{(i)}}{N}\right\}, \ldots, \\ & Y_1(N-1)\exp\left\{-\frac{j2\pi(N-1)\Delta_\tau^{(i)}}{N}\right\}\Bigg]^T, \\ & i = 1,\ldots,P_2, \end{aligned} \tag{49}$$

$$\boldsymbol{\beta} = \left[\frac{\alpha_{2;1}}{\alpha_{1;1}}, \frac{\alpha_{2;2}}{\alpha_{1;1}}, \ldots, \frac{\alpha_{2;P_2}}{\alpha_{1;1}}\right]^T, \tag{50}$$

and

$$\boldsymbol{\Delta}_\tau = \left[\Delta_\tau^{(1)}, \Delta_\tau^{(2)}, \ldots, \Delta_\tau^{(P_2)}\right]^T, \tag{51}$$

is the vector of the TDOAs of interest. $\boldsymbol{W}_p$ is a complex white Gaussian noise vector containing the samples $\{W_p(k)\}_{k=0}^{N-1}$. Considering these notations, it turns out that the estimation of the TDOAs can be performed using the same algorithm developed above for the active system. We only have to substitute the vector $\boldsymbol{\phi}_a(\tau)$ by $\boldsymbol{\phi}_p(\Delta_\tau)$ and $X(k)$ by $Y_1(k)$ in the expression of the periodogram in (24). The remaining steps follow in the same way.

Now we rediscuss the estimation problem when P_1 is different from 1. The problem then consists of estimating $P_1 \times P_2$ different parameters. To that end, we refer again to the results of the active case. $P_1 \times P_2$ values are generated according to $\overline{g}_{\rho_1}(.)$ by substituting $X(k)$ and $Y(k)$ in the expression of $I(.)$ by $Y_1(.)$ and $Y_2(.)$, respectively. Then, the generated values are classified from the smallest to the highest and organized as follows:

$$\boldsymbol{\Delta}_{\tau,k} = \left[\boldsymbol{\Delta}_{\tau,k}^{(1)}, \boldsymbol{\Delta}_{\tau,k}^{(2)}, \ldots, \boldsymbol{\Delta}_{\tau,k}^{(P_1)}\right], \tag{52}$$

where each vector $\{\boldsymbol{\Delta}_{\tau,k}^{(i)}\}_{i=1}^{P_1}$ is formed from P_2 TDOAs. The final step consists of evaluating the following means as in (29):

$$\Delta_{\tau,k}^{m,n} = \frac{1}{2\pi T}\angle\frac{1}{R}\sum_{r=1}^{R} F\left(\boldsymbol{\Delta}_{\tau,k}^{(m)}\right)\exp\left\{j2\pi\frac{\boldsymbol{\Delta}_{\tau,k}^{(m)}(n)}{T}\right\}. \tag{53}$$

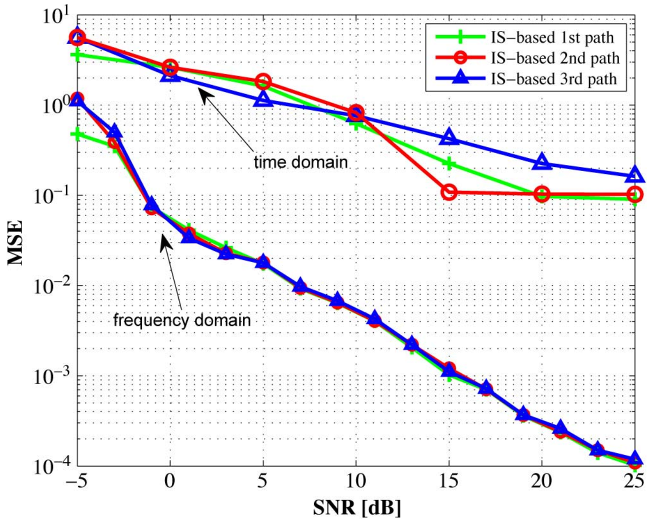


Fig. 3. Estimation performance of the IS-based algorithm in the time and frequency domains.

V. Simulation Results

To properly assess the performance of our new IS-based approach, we compare the performance of the proposed IS-based method to the *expectation maximization* (EM) algorithm [6] as one representative example of the iterative implementations of the ML criterion and to the MUSIC-type algorithm proposed in [5] as one representative example of suboptimal subspace-based methods. We also compare our solution to the ARS technique [20], a numerical algorithm developed to optimize non-linear functions, by applying it to the likelihood function. The estimation error of the different estimators is also compared to the Cramér-Rao lower bound (CRLB) which reflects the theoretical achievable performance taken as a benchmark for all the considered algorithms. In all simulations, the transmitted pulse is a chirp signal which is widely used in radar and sonar applications. The number of samples is set to $N = 70$. We consider 3 propagation paths with closely-spaced delays $[3T_s, 6T_s, 8T_s]$. The multipath gain is assumed to be equal for the three paths and we fix R to 1000. Note that R affects the accuracy of the proposed estimator. Its value can also be freely chosen as high as possible and it is clear that a very large value of R results in a consistent estimate of the mean value computed in (29) as well-known from estimation theory. But, by increasing R, we are actually trading increased complexity for enhanced estimation accuracy. After extensive simulations, we found that a relatively small value of $R = 1000$ yields very satisfactory trade-off between performance and complexity, since the estimator reaches the CRLB over a wide SNR range. The SNR is defined as $\mathrm{SNR} = \frac{\boldsymbol{X}^H\boldsymbol{X}}{N_0}$.

To further illustrate previous discussion in Section IV-A on the net advantages of the new frequency-domain formulation of the IS-based algorithm against the alternative time-domain in [15], we compare the mean square error (MSE) performance of both resulting implementations. As expected there, essentially from the accuracy or inaccuracy of the frequency- and time-domain diagonal-matrix approximations of (21) and (36), respectively (see also Fig. 1), Fig. 3 clearly shows that the performance of the time-domain estimator would degrade considerably compared to what can be achieved by the proposed frequency-domain estimator.

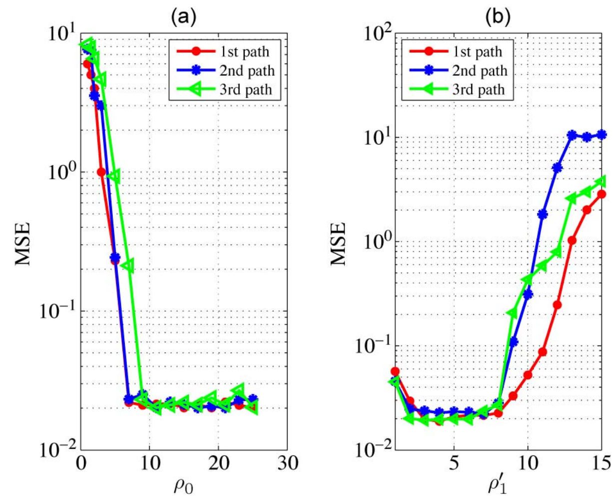


Fig. 4. Estimation performance vs. (a) ρ_0 and (b) ρ'_1.

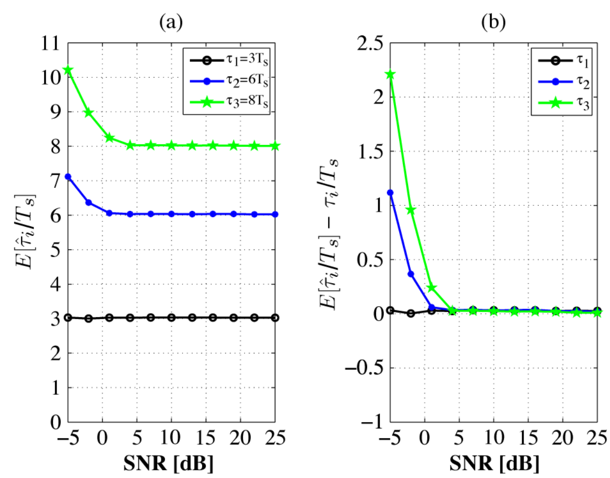


Fig. 5. Estimation bias of the IS-based algorithms vs. SNR (a) $E\left\{\frac{\hat{\tau}_i}{T_s}\right\}$ and (b) $E\left\{\frac{\hat{\tau}_i}{T_s}\right\} - \frac{\tau_i}{T_s}$.

As mentioned in Section IV, the algorithm is sensitive to the choice of ρ_0 and ρ'_1. To study the effect of each parameter on the estimation performance, we fix one of them and vary the other. It is seen from Fig. 4(a) that there is no dependence on the value of ρ_0 as soon as ρ_0 is larger than ρ'_1. On the other hand, as illustrated in Fig. 4(b), ρ'_1 affects seriously the estimation performance of the new IS-based algorithm. As already mentioned, small values of ρ'_1 may not reduce the effect of the additive noise involved in $\overline{g}_{\rho'_1}(.)$, while too large values reduce the desired lobes revealing the actual delays in $\overline{g}_{\rho'_1}(.)$ thereby preventing their generation. Therefore, an appropriate choice of ρ'_1 is necessary in order to obtain near-optimal performance. We see from Fig. 4 that for ρ'_1 taking values between 2 and 7, the performance is almost the same, and thus the optimal value of ρ'_1 can be freely selected from this relatively large range. In the following simulations, ρ_0 and ρ'_1 are set, respectively, to 15 and 6. For a non-symmetric likelihood function, the evaluation of the mean using (29) may introduce a bias in the estimation. But it can be shown that this bias can be significantly reduced by increasing the design parameter ρ_0. This property can be easily inferred from the proof of Pincus's theorem as originally introduced in [11]. Another remarkable point is, however, that the bias is larger for the delays that are located close to the edge

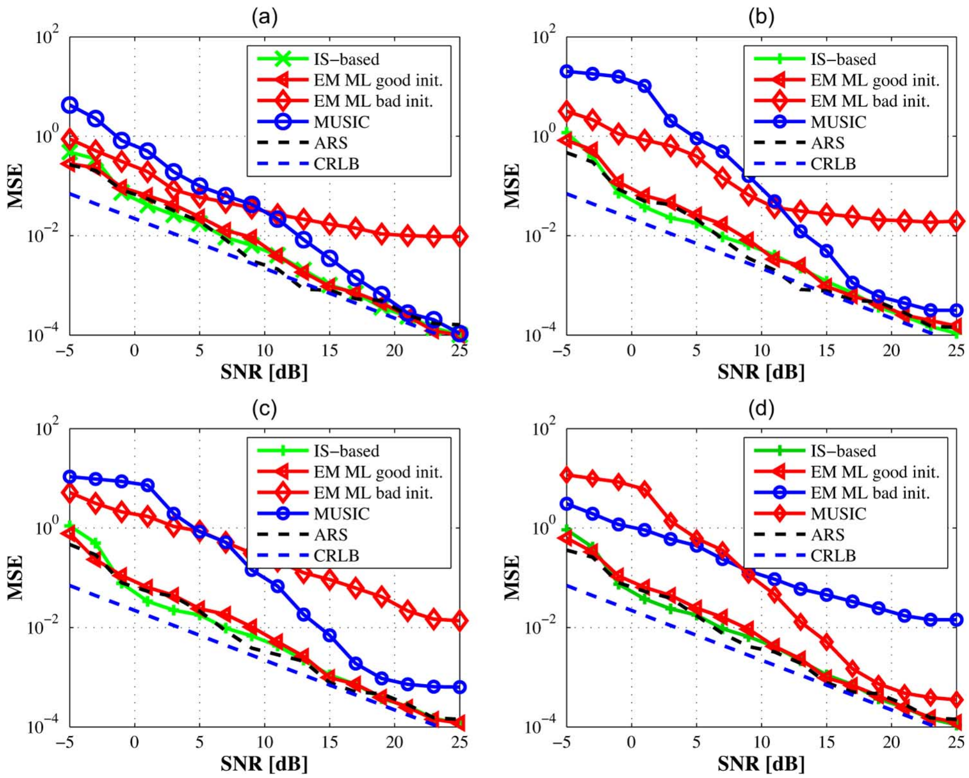


Fig. 6. Estimation performance of the IS-based, ARS-based, EM ML and the MUSIC-type algorithms vs. SNR in an active system for (a) 1st path, (b) 2nd path, (c) 3rd path, and (d) average over paths.

of the interval J (that is assumed to contain all the delays). The use of the circular mean in (29) eliminates indeed the bias that is due to edge effects. We corroborate our claims by computer simulations in Fig. 5 depicting the mean and the estimated bias associated to our estimator. This figure shows clearly that the bias is almost equal to zero for the three delays. Now returning to the comparison of the different estimators, we recall that the EM and IS-based algorithms are two different implementations of the ML criterion. They are hence expected to exhibit the same performance since they both try to maximize the same objective function. Yet, it should be kept in mind that the IS-based and MUSIC-type algorithms do not require any initialization while the EM algorithm is iterative in nature. Consequently, we consider for the EM algorithm two scenarios in which the initial values are selected as random variables, centered at the real time delays and having a variance of $4T_s^2$ and $10T_s^2$ reflecting, respectively, relatively accurate and less accurate initializations. Fig. 6 depicts the performance of these techniques. As expected, the three ML-based estimators perform better than the MUSIC-type estimator. However, for less accurate initialization, the performance of the EM algorithm deteriorates considerably over the entire SNR range. On the other hand, the ARS presents a good performance. But its computation cost remains very high since simulations reveal that it starts converging after 5000 realizations, 5 times more than required by the IS-based algorithm $(R = 1000)$. We see also that while the MUSIC-type technique approaches the CRLB only as far as the SNR is sufficiently high; the proposed algorithm performs close to the CRLB over the entire SNR range. This is hardly surprising since the IS-based estimator is a far more accurate implementation of the ML criterion. Same conclusions hold in the passive case whose performance results for all considered techniques are depicted in Fig. 7 for $P_1 = 1$ and $P_2 = 3$.

In Table I, we assess the complexity of the proposed IS-based estimator and compare it to the other algorithms in terms of computational intensity. Complexity stands here for the order of the number of complex-valued multiply and add operations of each technique based on its algorithmic description. In this table N_{iter} stands for the number of iterations for the EM and ARS algorithms and M is a parameter of the algorithm introduced in [5]. The third column of this table shows the ratio of complexities between all the techniques by taking our new IS-based technique as a reference (its number of operations in the denominator). The *grid search* algorithm is included in the table for complexity comparisons only and is not included in the simulations. It is seen that our new IS-based algorithm provides the best trade-off between complexity and accuracy. In fact, the EM-based method is slightly less complex but it exhibits serious performance degradation when it is not accurately initialized.

So far, comparisons have been performed vs. the SNR. To study the resolution power of the different estimators, we consider two propagations paths and vary the delay separation $\Delta\tau = \tau_1 - \tau_2$ at an SNR value of 10 dB. The results are shown in Fig. 8 for an active system. Clearly, as the difference between the delays gets smaller, the estimate becomes less accurate for all three methods. Yet the three ML-based estimators still perform better than the MUSIC-type algorithm. For well spaced delays, all the methods perform the same. Same results hold for the passive system but the simulations were not included for the sake of conciseness. Another important point to study is the effect of the signal bandwidth on the estimation performance. In fact, since all the derivations are made in the

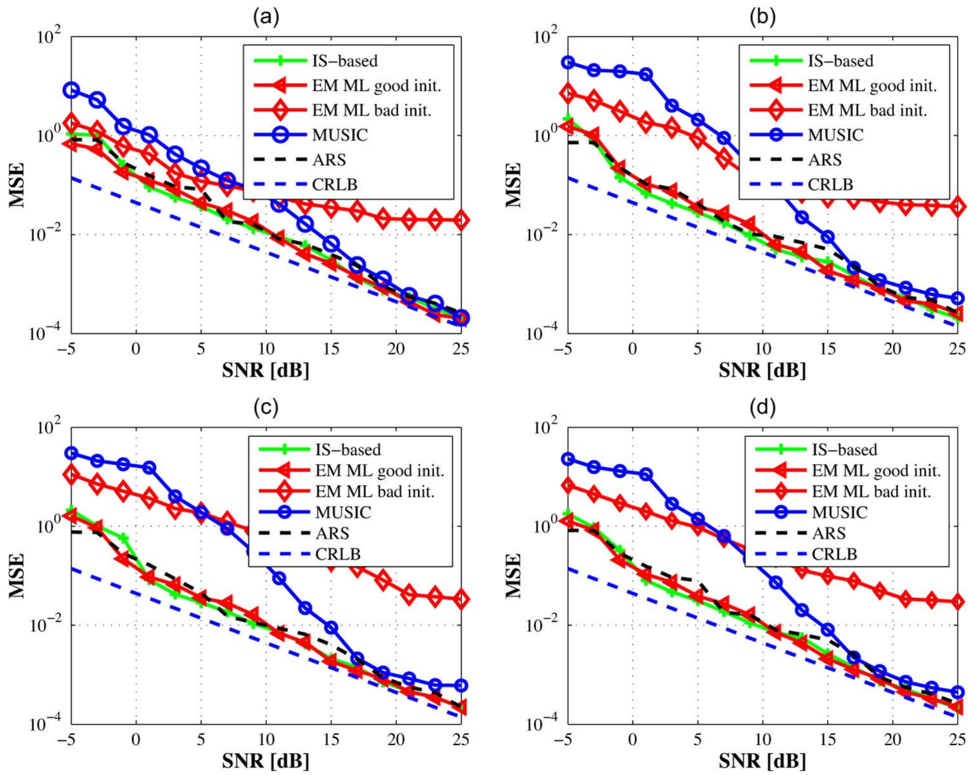


Fig. 7. Estimation performance of the IS-based, ARS-based, EM ML and the MUSIC-type algorithms vs. SNR in a passive system for (a) 1st path, (b) 2nd path, (c) 3rd path, and (d) average over paths.

TABLE I
COMPLEXITY ASSESSMENT OF THE ALGORITHMS

Algorithm	Complexity	Complexity ratio
IS-based	$O(2N\log(N) + 2NK + 4K + R(P^2N + P^3 + 2NP + PN + N) + P)$	1
Grid search	$O((P^2N + P^3 + 2NP + NP + N)K^P)$	10400
ARS	$O((P^2N + P^3 + 2NP + NP + N)N_{iter})$	3.452
EM	$O(P(K^2 + 2K)N_{iter})$	0.786
MUSIC-type	$O(N^2 + N(2N - 1) + 3M^3 + (2(M^2 + M(M - P)) + M)N + N(N - 1))$	1.109

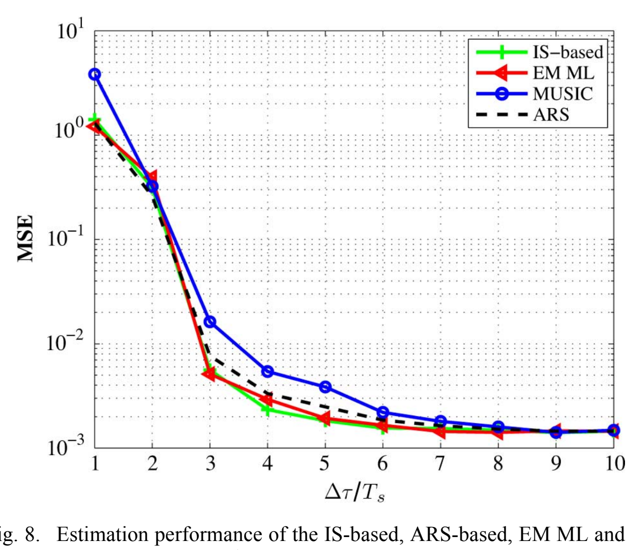


Fig. 8. Estimation performance of the IS-based, ARS-based, EM ML and the MUSIC-type algorithms vs. $\Delta\tau$ in an active system.

frequency-domain, the signal bandwidth (defined in the given example here as the difference between the higher and the lower frequency in the chirp signal) is expected to have an impact on the estimation procedure. Therefore, we compare in Fig. 9 the four estimators under different normalized signal bandwidth values (normalized by the sampling frequency i.e., W_bT_s where W_b is the signal bandwidth). Clearly, the proposed method outperforms the MUSIC-type algorithm over the entire bandwidth range, although the gap between the two methods decreases as the bandwidth increases while the performances of the IS-based and the ARS technique are almost the same. Note that the EM algorithm is also less sensitive to bandwidth variations. Same results hold for the passive system but the simulations were not included for the sake of conciseness.

Now we consider the case of time-varying channels. While the proposed method is primarily developed under the assumption of constant path gains, we verify through simulations that, in a Rayleigh-fading channel, it is also robust to time variations and that the ML-based estimators, namely the IS-based and the ARS algorithms, outperform MUSIC-type methods over relatively low Doppler frequency. Nonetheless, from Fig. 10 the performances of the two estimators, in an active system, degrade considerably as the normalized Doppler (i.e., f_DT_s) increases. In fact, the time variations of the channel coefficients

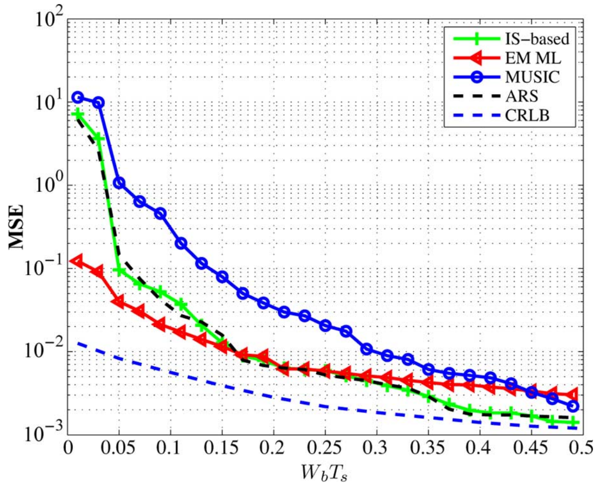


Fig. 9. Estimation performance of the IS-based, ARS-based, EM ML and the MUSIC-type algorithms vs. $W_b T_s$ at $\mathrm{SNR} = 10$ dB in an active system.

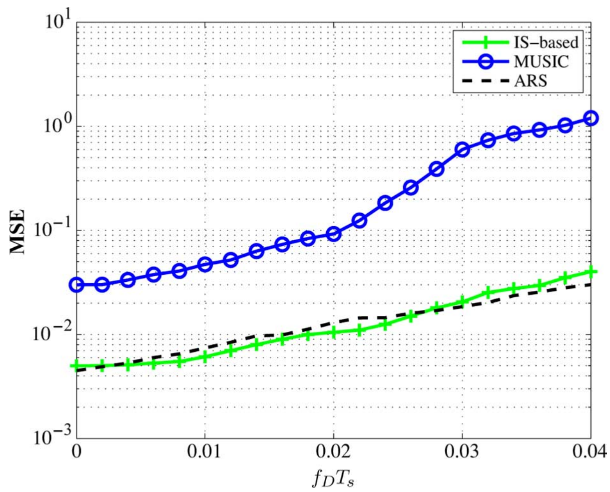


Fig. 10. Estimation performance of the IS-based, ARS-based and the MUSIC-type algorithms vs. $f_D T_s$ at $\mathrm{SNR} = 10$ dB in an active system.

are not taken into account when developing these algorithms, and it was shown in [22] that, in this case, the estimates become necessarily biased.[8] Note, in this case, that we are no longer able to obtain the estimates of the channel coefficients using (9). It is for this reason that the EM algorithm was omitted in this scenario since it is based, at each iteration, on an estimate of $\boldsymbol{\alpha}$, which cannot be performed for time varying channels.

VI. Conclusion

In this paper, we developed a new implementation of the ML-based estimator for multiple time delays based on the concept of importance sampling (IS). We considered the two cases of active and passive systems. The new algorithm is far less expensive in terms of computational complexity than the traditional multidimensional grid search method. Moreover, unlike the iterative methods, the IS-based algorithm does not suffer from initialization drawbacks. It performs well over the entire SNR range since its convergence to the global maximum of the likelihood function is guaranteed. In addition, it avoids the computational burden of the eigen-decomposition operation that is widely encountered in classical subspace-based techniques in multiple parameters estimation. On the other hand, the ARS algorithm performs the same as the IS-based algorithm. This could be expected since both algorithms return the global maximum of the likelihood function without depending on initialization. However, the ARS algorithm requires a higher number of realizations and thus results in an increase in the computational cost. In practice, the appropriate choice of the parameters ρ_0 and ρ_1', a critical issue for the IS-based algorithm, can be performed to further optimize the estimation performance.

[8]In [22], the effect of the time varying envelope has been treated in the case of frequency estimation with the MUSIC and ESPRIT algorithms.

Appendix
Method to Generate the Vector $\boldsymbol{\tau}$

In this appendix, we present some practical hints to easily generate a single realization of the vector $\boldsymbol{\tau}$.

— Define ξ as a discrete representation of the interval $[0, T]$ (i.e., $\xi = 0 : \frac{1}{s:T}$ with $\frac{1}{s}$ being a given step for some s).
— Generate τ_1 according to $\overline{g}_{\rho_1'}(.)$ using the inverse probability integration method. To do so, consider a random variables u uniformly distributed over [0, 1]. Then find $\tau_1 = \arg\max_{x\in\xi} |G(x) - u|$, where $G(x)$ if the cumulative distribution function associated to $\overline{g}_{\rho_1'}(.)$ (for more details, see [15]).
— Then eliminate the generated value τ_1 from ξ so that it cannot be generated again.
— Repeat the last two steps $P - 1$ times to generate $\tau_2, \tau_3, \ldots, \tau_P$ and obtain one realization of the P-dimensional vector $\boldsymbol{\tau}$.

References

[1] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, “A maximum likelihood time delay estimator using importance sampling,” presented at the Proc. IEEE GLOBECOM, Houston, TX, Dec. 5–9, 2011.
[2] M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed. New York: McGraw-Hill, 2001.
[3] G. C. Carter, *Coherence and Time Delay Estimation: An Applied Tutorial for Research, Development, Test and Evaluation Engineers*. New York: IEEE Press, 1993.
[4] M. C. Vanderveen, A.-J. Van der Veen, and A. Paulraj, “Estimation of multipath parameters in wireless communications,” *IEEE Trans. Signal Process.*, vol. 46, pp. 682–690, Mar. 1998.
[5] F.-X. Ge, D. Shen, Y. Peng, and V. O. K. Li, “Super-resolution time delay estimation in multipath environments,” *IEEE Trans. Circuits Syst.*, vol. 54, no. 9, pp. 1977–1986, Sep. 2007.
[6] M. Feder and E. Weinstein, “Parameter estimation of superimposed signals using the EM algorithm,” *IEEE Trans. Acoust., Speech. Signal Process.*, vol. 36, pp. 477–489, Apr. 1988.
[7] P. J. Hahn, V. J. Mathews, and T. D. Tran, “Adaptive realization of a maximum likelihood time delay estimator,” in *Proc. Conf. Acoust., Speech, Signal Process.*, May 1996, vol. 6, pp. 3121–3124.
[8] K. Gedalyahu and Y. C. Eldar, “Time delay estimation from low rate samples: A union of subspaces approach,” *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3017–3031, Jun. 2010.
[9] M. Pallas and G. Jourdain, “Active high resolution time delay estimation for large BT signals,” *IEEE Trans. Signal Process.*, vol. 39, no. 4, pp. 781–787, Apr. 1991.
[10] A. M. Bruckstein, T. J. Shan, and T. Kailath, “The resolution of overlapping echoes,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 6, pp. 1357–1367, Dec. 1985.
[11] M. Pincus, “A closed form solution for certain programming problems,” *Oper. Res.*, pp. 690–694, 1962.
[12] A. F. M. Smith and A. Gelfand, “Bayesian statistics without tears: A sampling-resampling framework,” *Amer. Statist.*, vol. 46, pp. 84–88, May 1992.
[13] H. Wang and S. Kay, “A maximum likelihood angle-Doppler etimator using importance sampling,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, pp. 610–622, Apr. 2010.

[14] H. Wang, S. Kay, and S. Saha, "An importance sampling maximum likelihood direction of arrival estimator," *IEEE Trans. Signal Process.*, vol. 56, pp. 5082–5092, Oct. 2008.
[15] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, "A non-data-aided maximum likelihood time delay estimator using importance sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4505–4515, Oct. 2011.
[16] L. Ng and Y. Bar-Shalom, "Multisensor multitarget time delay vector estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, pp. 669–677, Aug. 1986.
[17] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-59, pp. 1081–1089, Oct. 1986.
[18] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo methods in practice*. New York: Springer, 2001.
[19] K. Sharman, "Maximum likelihood parameter estimation by simulated annealing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 1988, pp. 2741–2744.
[20] M. J. Apple, R. LaBarre, and D. Radulovic, "On accelerated random search," *SIAM J. Optim.*, vol. 14, pp. 708–731, 2003.
[21] K. V. Mardia, *Statistics of Directional Data*. New York: Academic, 1972.
[22] O. Besson and P. Stoica, "Analysis of MUSIC and ESPRIT frequency estimates for sinusoidal signals with lowpass envelopes," *IEEE Trans. Signal Process.*, vol. 44, no. 9, pp. 2359–2364, Sep. 1996.

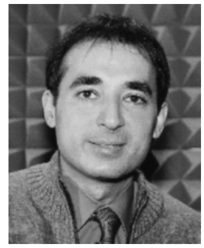
Ahmed Masmoudi was born in Ariana, Tunisia, on February 10, 1987. He received the Diplôme d'Ingénieur degree in telecommunication from the Ecole Supérieure des Communications de Tunis-Sup'Com (Higher School of Communication of Tunis), Tunisia, in 2010 and the M.Sc. degree from the Institut National de la Recherche Scientifique-Energie, Matériaux, et Télécommunications (INRS-EMT), Université du Québec, Montréal, QC, Canada, in 2012.

He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada. His research activities include signal processing and parameters estimation for wireless communication.

Faouzi Bellili was born in Sbeitla, Kasserine, Tunisia, on June 16, 1983. He received the Diplôme d'Ingénieur degree in signals and systems (with Hons.) from the Tunisia Polytechnic School in 2007 and the M.Sc. degree, with exceptional grade, at the Institut National de la Recherche Scientifique-Energie, Matériaux, et Télécommunications (INRS-EMT), Université du Québec, Montréal, QC, Canada, in 2009.

He is currently working toward the Ph.D. degree at the INRS-EMT. His research focuses on statistical signal processing and array processing with an emphasis on parameters estimation for wireless communications. During his M.Sc. studies, he authored/coauthored six international journal papers and more than ten international conference papers.

Mr. Bellili was selected by the INRS as its candidate for the 2009–2010 competition of the very prestigious Vanier Canada Graduate Scholarships program. He also received the Academic Gold Medal of the Governor General of Canada for the year 2009–2010 and the Excellence Grant of the Director General of INRS for the year 2009–2010. He also received the award of the best M.Sc. thesis of INRS-EMT for the year 2009–2010 and twice—for both the M.Sc. and Ph.D. programs—the National Grant of Excellence from the Tunisian Government. He was also rewarded in 2011 the Merit Scholarship for Foreign Students from the Ministère de l'Education, du Loisir et du Sport (MELS) of Québec, Canada. He serves regularly as a reviewer for many international scientific journals and conferences.



Sofiène Affes (S'94–M'95–SM'04) received the Diplôme d'Ingénieur degree in electrical engineering in 1992, and the Ph.D. degree (with Hons.) in signal processing in 1995, both from the Ecole Nationale Supérieure des Tétécommunications (ENST), Paris, France.

He has since been with the INRS-EMT, University of Quebec, Montreal, QC, Canada, as a Research Associate from 1995 until 1997, then as an Assistant Professor until 2000. Currently, he is an Associate Professor in the Wireless Communications Group. His research interests are in wireless communications, statistical signal and array processing, adaptive space-time processing, and MIMO. From 1998 to 2002, he has been leading the radio design and signal processing activities of the Bell/Nortel/NSERC Industrial Research Chair in Personal Communications at INRS-EMT, Montreal, QC, Canada. Since 2004, he has been actively involved in major projects in wireless communication of PROMPT (Partnerships for Research on Microelectronics, Photonics and Telecommunications).

Prof. Affes was the corecipient of the 2002 Prize for Research Excellence of INRS. He currently holds a Canada Research Chair in Wireless Communications and a Discovery Accelerator Supplement Award from NSERC (Natural Sciences & Engineering Research Council of Canada). In 2006, he served as a General Co-Chair of the IEEE Vehicular Technology Conference (VTC) 2006-Fall, Montreal, QC, Canada. In 2008, he received from the IEEE Vehicular Technology Society the IEEE VTC Chair Recognition Award for exemplary contributions to the success of IEEE VTC. He currently serves as a member of the Editorial Board of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the Wiley *Journal on Wireless Communications and Mobile Computing*.

Alex Stéphenne (S'94–M'95–SM'04) was born in Quebec, Canada, on May 8, 1969. He received the B.Eng. degree in electrical engineering from McGill University, Montreal, QC, Canada, in 1992 and the M.Sc. degree and Ph.D. degree in telecommunications from INRS-Télécommunications, Université du Québec, Montreal, QC, Canada, in 1994 and 2000, respectively.

In 1999, he joined SITA, Inc., Montreal, QC, Canada, where he worked on the design of remote management strategies for the computer systems of airline companies. In 2000, he became a DSP Design Specialist for Dataradio, Inc., Montreal, QC, Canada, a company specializing in the design and manufacturing of advanced wireless data products and systems for mission critical applications. In January 2001, he joined Ericsson and worked for over two years in Sweden, where he was responsible for the design of baseband algorithms for WCDMA commercial base station receivers. From June 2003 to December 2008, he was still working for Ericsson, but was based in Montreal, where he was a researcher focusing on issues related to the physical layer of wireless communication systems. Since 2004, he has also been an Adjunct Professor at INRS, where he has been continuously supervising the research activities of multiple students. His current research interests include Coordinated Multi-Point (CoMP) transmission and reception, Inter-Cell Interference Coordination (ICIC) and mitigation techniques in Heterogeneous Networks (HetNets), wireless channel modeling/characterization/estimation, statistical signal processing, array processing, and adaptive filtering for wireless telecom.

Dr. Stéphenne is a member of the organizing committee and a Co-Chair of the Technical Program Committee (TPC) for the 2012-Fall IEEE Vehicular Technology Conference (VTC'12-Fall) in Quebec City, Canada. He has served as a Co-Chair for the Multiple Antenna Systems and Space-Time Processing track for VTC'08-Fall in Calgary, Canada, and as a Co-Chair of the TPC for VTC'06-Fall in Montreal, Canada.